

Express Mail No. EL576314317US

Filing Date April 27, 2001

ATTORNEY DOCKET NO: 21419/91512  
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE  
UTILITY PATENT APPLICATION

INVENTORS: BRAD BOSWORTH  
PETER VÖGELI

TITLE: METHODS TO IDENTIFY SWINE GENETICALLY  
RESISTANT TO F18 *E. COLI* ASSOCIATED  
DISEASES

ATTORNEY: ALICE O. MARTIN  
BARNES & THORNBURG  
2600 CHASE PLAZA  
10 SOUTH LASALLE  
CHICAGO, ILLINOIS 60603

-1-

**METHODS TO IDENTIFY SWINE GENETICALLY  
RESISTANT TO F18 *E. COLI* ASSOCIATED DISEASES**

Inventors: Brad Bosworth and Peter Vögeli

5           This application claims priority to U.S. Provisional Application 60/047,181  
filed May 10, 1997, now abandoned; PCT/US98/10318, filed May 20, 1998; and is a  
divisional of U.S. Serial No.09/443,766 filed November 19, 1999.

          Compositions and non-invasive methods are provided for the identification of  
swine genetically resistant to *E. coli* bacteria supplied with fimbriae F18. DNA  
10   polymorphisms in the swine alpha (1,2) fucosyltransferase (*FUT1*) gene were  
identified that differentiate resistant from susceptible swine and provide a diagnostic  
test useful for swine breeders.

          A major problem in breeding swine is to keep them disease-free. Intestinal  
disorders postweaning are a particular problem. A limited number of serotypes of  
15   toxigenic *Escherichia (E.) Coli* strains are the causative agents of oedema disease and  
postweaning diarrhea in swine which induce serious economic losses, especially  
among piglets aged 4 to 2 weeks, in swine breeding farms all over the world. The  
typical symptoms of oedema disease are neurological signs such as ataxia,  
convulsions and paralysis. At post mortem examination, oedema is typically present  
20   at characteristic sites such as eyelids and forehead, stomach wall and mesocolon. The  
diseases are caused by Shiga-like toxin-II variant and enterotoxins LT, Sta, Stb  
respectively, produced by *E. coli* that colonize the surface of the small intestine  
without effecting major morphological changes of the enterocytes (cells in the  
intestine). Certain types of bacterial *E. coli* strains, F18, F4 and K88 are major lethal  
25   villains in this regard. "Oedema disease of pigs is an enterotoxaemia characterized by  
generalized vascular damage. The latter is caused by a toxin, Shiga-like toxin II  
variant, produced by certain strains of *E. coli*" (Bertschinger *et al.*, 1993). The *E. coli*  
are distinguished by their pili types, a group of adhesive fimbriae that are related are  
designated *e.g.*, K88 or F18 (Vögeli *et al.*, 1997).

30   Not all swine succumb to *E. coli* infections. Colonization depends on  
adherence of the bacteria to the enterocytes which is mediated by the bacterial  
fimbriae designated *e.g.*, K88 or F18. Susceptibility to adhesion, *i.e.* expression of

receptors in swine for binding the fimbriae, has been shown to be genetically controlled by the host and is inherited as a dominant trait with, in the case of F18, *B* being the susceptibility allele and *b* the resistance allele (Vögeli *et al.*, 1996; Meijerink *et al.*, 1996). The genetic locus for this *E. coli* F18-receptor (*ECF18R*) has been mapped to porcine chromosome 6 (SSC6), based on its close genetic linkage to the *S* locus and other loci of the halothane (HAL) linkage group on chromosome 6. The receptor for K88 *E. coli* is on chromosome 13.

The mechanism for resistance appears to be that intestinal borders in resistant animals are not colonized by *E. coli*, i.e., the bacteria do not adhere to intestinal walls of resistant swine. Glycoprotein receptors in the brush border membrane of the intestine were shown to be responsible for the differences between adhesive and non-adhesive phenotypes related to some *E. coli*, therefore, the genotype of the host swine determines resistance. The fimbriated bacteria also have been studied (WO 9413811).

Current methods of identifying swine that are resistant to F18 *E. coli* associated diseases are either to 1) collect intestinal samples from swine at slaughter and perform the microscopic adhesion test, 2) challenge the animals with virulent *E. coli* ("colonization test"), or 3) perform blood typing of the A-O(S) blood group system. The first two methods are not practical for identifying resistant animals for use as breeding stock. Although the blood typing method does identify resistant animals, the test is unable to determine whether susceptible animals are homozygous or heterozygous for susceptibility. Knowledge of the genotype of animals with regard to these alleles (conditions of a gene) is essential to develop a successful breeding program. The purpose of the breeding program is to produce swine that are resistant to F18 *E. coli* associated diseases that decimate stock post-weaning.

In one publication the authors stated, in reference to oedema disease in swine, that "Searches are underway for appropriate genetic markers. . ." (Bertschinger *et al.*, 1993, page 87) and, citing Walters and Sellwood, 1982:

Breeding resistant swine is an attractive method for prevention of diseases for which an effective prophylaxis is not available. The feasibility of this approach will depend on the prevalence of the gene(s) encoding resistance in the pig population, improved methods for the detection of resistant pigs, and absence of negative genetic traits co-selected with this resistance.

A genetic "marker" locus is a coding or non-coding locus that is close to a genetic locus of interest, but is not necessarily the locus itself. Detectable phenotypes include continuous or discontinuous traits, *e.g.* restriction length fragment polymorphisms, production traits, bacterial adhesion traits, colorimetric or enzymatic reactions, and antibiotic resistance. The *S* locus controls expression of the A and O blood group antigens. Swine homozygous recessive at the *S* locus do not express either A or O blood group antigens. A similar condition exists in humans and is due to mutations in the alpha (1,2) fucosyltransferase gene which encodes the human blood group H (Kelly *et al.*, 1994; *see also* WO 9628967). The porcine alpha (1,2) fucosyltransferase gene of swine has recently been sequenced (Cohney *et al.*, 1996). This gene is very likely the gene present at the *S* locus in swine.

The blood group *H* and *Se* loci have been mapped genetically and physically to human chromosome 19q13.3. This region is evolutionarily conserved, containing genes homologous to the HAL linkage group of genes in pigs. The blood group H encoding gene is the so called *FUT1* whereas the *Se* gene is equivalent to the *FUT2* gene. *FUT1* determines H antigen expression in the erythroid cell lineage, whereas *FUT2* regulates expression of the H antigen in the secretory epithelia and saliva. Conservation of the *FUT1* gene has been shown in lower mammals such as rat and rabbit, and mRNA expression has been shown in rabbit brain tissue and rat colon. In all these species two types of alpha (1,2) fucosyltransferase genes have been reported which are structurally very similar to the human *FUT1* and *FUT2* genes, but in particular the *FUT1* homologous genes show a species specific expression pattern. In humans the *FUT1* gene is responsible for synthesis of H antigens in the precursors of erythrocytes. However, in pigs erythrocytes passively absorb H-like antigens from the serum, as is the case for the human *Lewis* antigens. In pigs all H-like antigens are related to exocrine secretory tissues, and expression of the *FUT2* (*Secretor*) gene is seen in secretory tissue of other animal species. Therefore, expression of the porcine A-O blood group determinants which cross-react with anti-human blood group H and A antibodies might be influenced by the *FUT2* gene.

Further information about blood groups and *E. coli* swine diseases include that carbohydrate structures of blood group antigens mediate the adhesion of some pathogenic microorganisms to host tissues, *e.g.* *Helicobacter pylori* adhere to Lewis<sup>b</sup>

blood group antigens, and *E. coli* causing urinary tract infections adhere to blood group *P* substance. Genes encoding glycosyltransferases that are responsible for the formation of the blood group specific carbohydrate structures, therefore, represent candidate genes for the control of bacterial colonization by the host. The localization of these genes is in the same chromosomal region as the locus responsible for adhesion/non-adhesion of F18 positive *E. coli* in the swine small intestine. Swine do not express blood group antigens A and O until after weaning, this is the same time that they become susceptible to disease caused by F18 *E. coli*.

New methods of diagnosis and treatment are needed for *E. coli* related intestinal diseases in swine. Detection of a genetic mutation was proposed as a diagnostic test for some swine disorders (malignant hypothermia) (Fujii *et al.*, 1991; U.S. Pat. No. 5,358,649), but polymorphic markers were not reported for diagnosis. Vaccines to develop resistance to *E. coli* colonization were described (U.S. Pat. No. 5,552,144; WO 8604604), but are unlikely to be a preferred method to prevent the *E. coli* disease because of difficulties in administering live vaccine orally to newborn swine, and because of regulatory restrictions. Antibiotics are available for treatment, but there is no successful prophylaxis.

#### **SUMMARY OF THE INVENTION**

The compositions and non-invasive methods of the present invention provide detection and elimination of swine that are susceptible to *E. coli* associated diseases. A non-invasive method for identifying a swine that is resistant to intestinal colonization by *E. coli* F18 includes the following steps: determining whether a genetic polymorphism associated with resistance to colonization is in a biological sample from the swine; and inferring that the swine is resistant if the swine is homozygous for the polymorphism.<sup>1</sup>

More particularly, the method is determining in a biological sample from the swine whether the nitrogen base at position 307 in the alpha (1,2) fucosyltransferase gene of the swine is only adenine or only guanine; and identifying the swine as resistant if the only nitrogen base at position 307 is adenine.

---

<sup>1</sup>

A polymorphism is a change in a nucleotide sequence that exists in a population due to mutation.

To determine whether a polymorphism is present in a biological sample, restriction fragment length polymorphisms are analyzed on a gel that separates them by molecular weight. Restriction endonucleases are enzymes that reproducibly cut nucleic acid molecules at specific sites, resulting in nucleic acid fragments of different molecular weights, depending on the location of the cuts.

The invention also relates to a method for breeding swine to be resistant to *E. coli* associated diseases by selecting for breeding swine that have a genetic polymorphism in the alpha (1,2) fucosyltransferase 1 gene that identifies them as swine that are resistant to *E. coli* related intestinal diseases; and breeding the selected swine.

An aspect of the invention is a DNA molecule which is polymorphic for the alpha (1,2) fucosyltransferase 1 gene in swine, in particular a sequence in accordance with FIG. 1. Other aspects of the invention are molecules with nucleotide sequences complementary to that of FIG. 1.

An aspect of the invention is an isolated DNA molecule with a substitution of adenine for guanine in position 307. This molecule may also bond a substitution of adenine for guanine in position 857. Other isolated DNA molecules of the present invention include those with a mutation at nucleotide position 229 of the sequence of FIG. 1, wherein the codon CTT is changed to TTT, encoding for the amino acid phenylalanine instead of leucine. A mutation at nucleotide position 714 is from GAT → GAC, but there is no accompanying amino acid substitution in the encoded product.

Polypeptides encoded by the DNA molecules of the present invention and having alpha (1,2) fucosyltransferase activity are also aspects of the invention.

A molecular assay for detecting *E. coli* F18 receptors in swine is to (a) isolate DNA from porcine nucleated cells; (b) amplify the DNA in a polymerase chain reaction (PCR) using oligonucleotides as primers which are complementary to a DNA sequence of the porcine alpha (1,2) fucosyltransferase gene 1; (c) perform a restriction enzyme digest with at least one restriction enzyme *e.g.*, CfoI; (d) separate the resulting fragments by gel electrophoresis; and (e) determine the respective numbers and lengths of fragments on the gel; and (f) determine from the numbers and length of fragments of F18, which receptors are present in the porcine cells. Use of the larger amplified fragments disclosed herein for restriction length polymorphism analysis

(RFLP), rather than smaller fragments, is less expensive because the DNA bands can be run on agarose gels of relatively low concentration. Also, to produce some of the fragments, only one restriction enzyme is needed for a constant restriction site adjacent to the variable diagnostic site.

- 5       A kit for detecting polymorphisms associated with *E. coli* F18 receptors uses oligonucleotides in separate containers which are complementary to a DNA sequence of the porcine alpha (1,2) fucosyltransferase gene 1 that distinguishes resistant from sensitive swine. The test can be performed on swine of any age.

- 10       The polymorphisms are also useful to develop drugs to treat swine that have *E. coli*-associated disease. A mutated form of porcine alpha 1,2 fucosyltransferase could interfere with the normal enzyme, preventing it from producing the intestinal receptor for F18.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

- 15       FIG. 1 shows the nucleotide sequence (*FUT1*) (below) and the predicted amino acid sequence (above) of the swine  $\alpha$  1  $\rightarrow$  2 fucosyltransferase polymorphism of the present invention using the *one-letter* amino acid code. The *solid double line* below the amino acid sequences (=) is the putative transmembrane region; the *dotted line* below the amino acid sequence shows three potential N-linked glycosylation sites (. . .).

- 20       □ is where an adenine (A) is substituted for guanine (G) in resistant swine.

\* Indicates the termination codon.

Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

- 25       **DESCRIPTION OF THE PREFERRED EMBODIMENT**

Molecular analysis of DNA polymorphisms associated with resistance of swine to *E. coli* associated diseases facilitated diagnostic assays to select resistant pigs for breeding. Resistant pigs differ from sensitive pigs at the *E. coli* F18 receptor locus as identified by the polymorphic markers of the present invention.

- 30       The present invention provides non-invasive methods and compositions to distinguish with a high level of sensitivity and specificity, swine that are genetically susceptible to diseases associated with F18 *E. coli* infection from resistant swine. A

DNA polymorphism in the swine alpha (1,2) fucosyltransferase (*FUT1*) gene was identified that differentiates resistant from susceptible swine. The polymorphism arose by a mutation (change) in a nucleotide sequence leading to a new allele. An allele is a condition of a gene. In a population there may be many alleles of a gene differing by nitrogen base substitutions, presumably caused by mutations in an ancestral DNA molecule. The coexistence in a population of more than one allele (sometimes referred to as a "variant") is called a genetic polymorphism. Loci at which more than one allele may exist as apparently stable components of a population, is a polymorphic locus. Usually, one of the polymorphic loci is at a low frequency in the population.

As determined from a biological sample, preferably blood, the resistant swine have a polymorphism in their genomes in which the only base detected at position 307 (see FIG. 1) in the nucleotide sequence is adenine, whereas the base in the same position in homozygous susceptible swine is guanine. Heterozygous swine will show both types of DNA and will be susceptible. The polymorphism is a variation of a porcine gene sequence (Cohney *et al.*, 1996).

Genetic linkage analysis was performed on families of swine and genetic associations between polymorphisms in *FUT1* and disease resistance in outbred swine were determined. According to the present invention, polymorphisms have been found in the alpha (1,2) fucosyltransferase 1 gene (*FUT1*). A polymorphism that has a single nucleotide base substitution at position 307 was used to establish a close linkage between the fucosyltransferase gene and the *S*-system, the *ECF18R* locus and other loci of the *HAL* linkage group.

The detection of the close linkage of the mutation at *FUT1* and *ECF18R* allowed a molecular test to be developed for the identification of *E. coli* F18 adhesion resistant, heterozygous (carrier) and homozygous susceptible pigs. This diagnostic test identifies, with high sensitivity and specificity, pigs that are susceptible to oedema disease and postweaning diarrhea. The incidence of the polymorphisms of the present invention differs among swine breeds. Vögeli, *et al.* (1997) presented frequencies of the M307 allele in 5 pig breeds from herds that were not related to one another. The availability of the diagnostic test for the polymorphism of the present invention provides breeders with the opportunity to effectively eliminate the *ECF18R*



susceptible allele from their swine herd, thereby eliminating a prerequisite for *E. coli* F18 bacterial adhesion causing oedema disease and postweaning diarrhea.

The present invention further includes nucleotide sequences that are variants of a sequence of the alpha (1,2) fucosyltransferase gene 1 representing the various polymorphisms at bp 307, and diagnostic molecular based kits to identify polymorphisms in the alpha (1,2) fucosyltransferase gene.

In order to obtain candidate genes for the *E. coli* F18 receptor locus (*ECF18R*) 5 cosmids and one genomic clone containing the gene were isolated containing the alpha (1,2) fucosyltransferase genes, *FUT1* and *FUT2* (Meijerink *et al.*, 1997), from a porcine genomic library. Mapping by fluorescence *in situ* hybridization placed all these clones in band q11 of porcine chromosome 6 (SSC6q11). Sequence analysis of the cosmids resulted in the characterization of (a) an open reading frame (ORF), 1098 base pairs in length, that is 82.3% identical to the human *FUT1* sequence, and (b) a second ORF, 1023 base pairs in length which is 85% identical to human *FUT2* sequence. The *FUT1* and *FUT2* loci therefore seem to be porcine equivalents of the human blood group *H* and the *Secretor* locus. Direct sequencing of the two ORF's in swine either susceptible or resistant to adhesion and colonization by F18 fimbriated *E. coli* (*ECF18R*) revealed two polymorphisms at base pair 307 (M307) and base pair 857 (M857) of the *FUT1* ORF. The nucleotide positions are numbered from ATG (encoding methionine). Analysis of these mutations in 34 matings of Landrace families with 221 progeny showed close linkage with the locus controlling resistance and susceptibility to *E. coli* F18 adhesion and colonization in the small intestine (*ECF18R*) and with the locus of the blood group inhibitor *S*. Therefore, the M307 mutation is a good marker for marker-assisted selection of *E. coli* F18 adhesion resistant animals. Another mutation at nucleotide position 229 was found leading to a polymorphism in which the codon encoding leucine (CTT) was changed to TTT (encoding phenylalanine). A mutation at position 714 (GAT → GAC) (encoding aspartic acid) did not produce an amino acid substitution. No polymorphisms were identified in *FUT2* that differentiated susceptible and resistant pigs.

### EXAMPLES

The following examples provide embodiments of the invention.

**Example 1: An Assay For Resistant Swine**

The polymorphisms of the present invention are easily identified using PCR-RFLP tests. One embodiment of the tests used a 160bp fragment of porcine alpha (1,2) fucosyltransferase 1 amplified using PCR with the following primers;  
 5'CCAACGCCTCCGATTCCTGT3' and 5'GTGCATGGCAGGCTGGATGA3'. Preferred PCR conditions for this embodiment are 25 cycles at the following times and temperatures: 94°C, 30 sec; 60°C, 45 sec; 72°C, 90 sec. The amplified DNA from resistant swine was digested by the restriction enzyme Hgal, but was not digested by the restriction enzyme HinPI. The amplified DNA from homozygous susceptible swine was digested by the restriction enzyme HinPI. The amplified DNA from heterozygous susceptible swine was partially digested by both enzymes.

Alternatively, DNA was isolated from porcine nucleated cells according to standard procedures. Direct sequencing of porcine *FUT1* and *FUT2* sequences and their flanking regions in animals of different *ECF18R* genotype (Bb, bb) resulted in the identification of two G -> A transitions at positions 307 and 857 (termed *M307* and *M857*, respectively) of the *FUT1* ORF. The *M307* transition eliminates a restriction site for CfoI. Amplification of DNA isolated from porcine nucleated cells was preformed according to standard procedures with primers P6 and P11 (3 min at 95°C, 30 cycles of 30 sec at 95°C, 30 sec at 56° C and 30 sec at 72°C, followed by a 7 min final extension at 72°C) followed by CfoI digestion and separation on a 3% agarose gel resulted in a restriction fragment length polymorphism (RFLP). Homozygous *M307<sup>AA</sup>* animals showed 2 bands. Homozygous *M307<sup>GG</sup>* animals showed 93-, 241- and 87bp fragments. Heterozygous animals showed all four fragments.

**Example 2: Sensitivity and Specificity Of An Assay Using Alpha (1,2) Fucosyltransferase In Detecting Swine Resistant to F18 *E. coli***

A study was conducted to determine the association between disease resistance and the polymorphism at position 307 of the *FUT1* gene. 183 weaned swine (ranging in ages 2-6 months) were obtained from six different breeding herds. Only one of these herds was known to contain resistant animals before the start of the study, and this herd is known to have a high incidence of porcine stress syndrome. The other 5 herds had no evidence of porcine stress syndrome, and the incidence of disease resistance was unknown. Swine from each herd were randomly selected, humanely euthanized and spleens and samples of small intestine were removed. DNA was extracted from splenic

tissue and used in a PCR-RFLP assay described in Example 1. Intestinal cells were purified by scraping the mucosal surface off the intestine, lysing the cells in a hypotonic EPTA solution and washing by centrifugation. The purified intestinal cell brush borders were incubated with F18 *E. coli*. This mixture was examined by phase contrast microscopy. This assay determined if swine were susceptible (intestinal samples had adhering bacteria) or resistant (intestinal samples had no adhering bacteria). The PCR-RFLP assay for the polymorphism correlated with the bacteria-intestinal cell binding assay in 53 of 53 resistant swine and 128 of 130 susceptible swine. Two swine that were determined susceptible using the bacteria-intestinal cell binding assay were incorrectly predicted to be resistant using the PCR-RFLP assay. Two of the six herds examined contained resistant pigs, while only one herd had porcine stress syndrome, demonstrating that the PCR-RFLP assay can identify disease resistant animals in animals that do not have porcine stress syndrome.

**Example 3: Localization of *FUT1* on Chromosome 6 (SSC6)**

Cosmids ETHs1, -s2, -s3, -s4 and -s6 were identified after screening of the cosmid library with a *FUT1* nucleotide probe obtained from porcine genomic DNA with primers P7 and P10 and were mapped by FISH and DISC-PCR to chromosome 6 in band q11.

**Example 4: Identification of the Porcine *FUT1* ORF**

Hybridizing KspI, EcoRI and KspI/EcoRI cosmid digests with radiolabelled porcine *FUT1* fragments P6-P11 and P7-P10 for Southern blot analysis revealed identical autoradiography signals for ETHs2, -s4 and -s6, whereas different signals were obtained from cosmids ETHs1 and -s3. From cosmid ETHs2 KspI, subclones 940 bp and 6.2 kb in length were isolated, corresponding to the estimated length of hybridizing KspI fragments on the Southern blot. The sequence results of both subclones were combined to yield a 1501 bp sequence, which was in agreement with results of direct sequencing of genomic PCR products. The 1501 bp sequence contains an open reading frame (ORF) of 1098 bp corresponding to the human *FUT1* ORF, with 82.3% nucleotide and 80.8% amino acid identity. The ORF encodes a polypeptide.

**Example 5: Identification of a Porcine *FUT2* and a Pseudogene *FUT1***

ETHs1 has one DNA fragment (2.7 kb) that hybridizes to *FUT1* sequences, whereas ETHs3 has two (2.7 kb and 8.2 kb). Subcloning and partial sequencing of the 2.7 kb EcoRI fragment of ETHs1 and -s3 confirmed that these two fragments are identical. The sequence is highly similar to the human *FUT2* but shows several changes

in the NH<sub>2</sub>- and -COOH terminal regions. These changes lead to frame shifts that are not compatible with a conserved ORF, therefore an assumption is that the sequence obtained from the 2.7 kb fragment represents a pseudogene (*FUT2P*). After subcloning of ETHs3 BamHI digests, the hybridizing sequences contained in the 8.2 kb EcoRI fragment were

5 identified. The sequence of the subclones obtained represents a 1023 bp ORF and is 85% identical at the nucleotide- and 83%- identical at the amino acid level to the human *FUT2* sequence. Many differences in the NH<sub>2</sub>- and -COOH terminal regions were observed between the porcine *FUT2* sequence and the *FUT2P* sequence derived from the 2.7 kb fragment. The predicted amino acid sequence corresponds to the partially determined

10 amino acid sequence of the porcine *Secretor* enzyme (Thurin and Blaszczyk-Thurin, 1995). The porcine *FUT1*, *FUT2*, and *FUTP* sequences obtained were submitted to GenBank and have accession numbers U70883, U70881 and U70882, respectively. The *FUT1* and *FUT2* genes have highly homologous sequences. This has to be considered in, for example, primer development. Furthermore, *FUT1* and *FUT2* enzyme activity

15 need to be differentiated in further studies.

**Example 6: Identification of M307 and M857 Mutations and  
Characterization of M307**

DNA was isolated from porcine nucleated cells according to standard procedures.

20 Direct sequencing of porcine *FUT1* and *FUT2* sequences and their flanking regions in animals of different *ECF18R* genotypes (*Bb*, *bb*) resulted in the identification of two G→A transitions at positions 307 and 857 (termed *M307* and *M857*, respectively) of the *FUT1* ORF. The *M307* transition eliminates a restriction site for the enzyme CfoI. Amplification of DNA isolated from porcine nucleated cells was performed according

25 to standard procedures with primers P6 and P11 (3 min at 95°C, 30 cycles of 30 sec at 95°C, 30 sec at 56°C and 30 sec at 72°C, followed by a 7 min final extension at 72°C) followed by CfoI digestion and separation on a 3% agarose gel resulted in a restriction fragment length polymorphism (RFLP). Homozygous *M307<sup>AA</sup>* animals showed 2 bands (93- and 328-bp fragments). Homozygous *M307<sup>GG</sup>* animals showed 87-, 93-, an 241-bp

30 fragments. Heterozygous animals showed all four fragments.

**Example 7: Characterization of Mutation M857**

The M857 mutation is a transition that eliminates an AciI site. Primer PBEST was designed to mismatch two additional AciI sites at positions 866 and 872. PCR with primers P7 and PBEST (3 min at 95°C, 30 cycles of 30 sec at 95°C, 30 sec at 56°C and

30 sec at 72°C, followed by a 7 min final extension at 72°C) followed by *AciI* digestion enables PCR-RFLP analysis on a 3% agarose gel. Homozygous *M857<sup>AA</sup>* animals show a 174 bp fragment while amplification products of *M857<sup>GG</sup>* animals show 136- and 38-bp fragments.

5 **Example 8: Genetic Mapping of the *FUT1* Gene**

- In Landrace swine families, recombination events between *M307* and the loci of the HAL linkage group (*S*, *ECF18R*, *RYR1*, *GPI*, *PGD*) revealed recombination fractions  $\Theta < 0.04$  (Table 2). The lodscores *Z* for the overall recombination fractions were between 24.5 and 50.6, showing strong evidence for linkage between these loci. These data allow
- 10 genetic mapping of the *FUT1* gene to the HAL linkage group in close proximity of *S* and *ECF18R* which are both influenced by *FUT1*. In experimental Landrace families, allelic association was found between *ECF18R* and *RYR1*. An excess of genotypes *RYR1<sup>IT</sup>* at position 1843 in *RYR1* (halothane susceptible genotype) was observed among pigs resistant to oedema disease and postweaning diarrhea (genotype *ECF18R<sup>bb</sup>*) (Table 3).
- 15 This allelic association is a result of linkage disequilibrium, that is, deviation of observed haplotype frequencies from expected haplotype frequencies under independent assortment of alleles. Therefore, linkage disequilibrium designates a non random association of alleles belonging to linked loci. Owing to low recombination rates however, no locus order could be determined as being significantly better than others.
- 20 **Example 9: Association of *M307<sup>A</sup>* With *ECF18<sup>b</sup>* And *M307<sup>G</sup>* With *ECF18R<sup>B</sup>***

- In Landrace (SL) and Large White (LW) parental pigs, *ECF18R<sup>b</sup>* (the oedema and postweaning diarrhea **resistance** allele) is 100% associated with *M307<sup>A</sup>*, and *ECF18R<sup>B</sup>* (the oedema and postweaning diarrhea **susceptibility** allele) is 100% associated with *M307<sup>G</sup>* (wherein A=adenine; G=guanine). In SL pigs 88% (30/34) of *S<sup>s</sup>* accounted for
- 25 all *ECF18R<sup>b</sup>* and *M307<sup>A</sup>* haplotypes, respectively. The corresponding values for both the *S<sup>s</sup>-ECF18R<sup>b</sup>* and *S<sup>s</sup>-M307<sup>A</sup>* haplotypes were 82% (9/11) in Large White pigs. In the experimental SL families, the occurrence of the *M857<sup>a</sup>* allele at the *FUT1* locus was low, and even absent, in LW pigs. Therefore, a significant gametic association was not observed between the alleles of *M857* and the alleles of the flanking genes. The G->A
- 30 transitions at positions *FUT1* 307 and *FUT1* 857 were found with variable frequencies also in Duroc, Hampshire and Pietrain pigs, making it likely that those transitions also occur in other pig breeds.

**Example 10: Distribution of *FUT1* Genotypes**

Table 4 shows that the distribution of *FUT1* genotypes at nucleotide position 307 among *ECF18R* types was significantly different from the expected ratio under a hypothesis that the two are independent. Of the 119 oedema disease and postweaning diarrhea resistant *ECF18R<sup>bb</sup>* animals, 118 were determined to have the genotype *M307<sup>AA</sup>* in the DNA-based test. One resistant animal had the genotype *M307<sup>AG</sup>*. Of the 131 susceptible pigs, 130 were *M307<sup>AG</sup>* or *M307<sup>GG</sup>*. One animal, susceptible to *E. coli* adhesion, was shown to be homozygous *M307<sup>AA</sup>* by the DNA-test. The data from this example and example 2, together with past studies suggested that the *FUT1* gene is the gene present at the *S* locus in swine and the *ECF18* locus. While 4 animals in this example and example 2 contradict this hypothesis, it is probable these animals were incorrectly phenotyped in regards to disease resistance/susceptibility.

**Example 11: Amino Acid Exchanges in Alpha (1,2) Fucosyltransferase**

The G → changes at bp +307 and bp +857 of the alpha (1,2) fucosyltransferase gene 1 results in a predicted amino acid substitution of threonine (neutral-polar) instead of alanine (neutral-nonpolar) and glutamine (neutral-polar) instead of arginine (basic), respectively which may have functional consequences in the encoded product. AC → T change at bp 229 results in an amino acid substitution of leucine (neutral-nonpolar) instead of phenylalanine (neutral-nonpolar).

**Table 1:** Sequences Of Forward-(F) And Reverse-(R) Primers And Their Relative Position to the Porcine *FUT1* and *FUT2* Start Codons<sup>2</sup>

	Primer name	Primer Sequence	Position
5	FUT1 P6 (R)	5'-CTTCAGCCAGGGCTCCTTTAAG-3'	+489
	FUT1 P7 (F)	5'-TTACCTCCAGCAGGCTATGGAC-3'	+720
	FUT1 P10 (R)	5'-TCCAGAGTGGAGACAAGTCTGC-3'	+1082
	FUT1 P11 (F)	5'-CTGCCTGAACGTCTATCAAGATC-3'	+69
	FUT1 P16 (F)	5'-AGAGTTTCCTCATGCCACAGG-3'	-90
10	FUT1 P18 (R)	5'-CTGCTACAGGACCACAGCATC-3'	+1203
	FUT1 PBEST (R)	5'-ACCAGCAGCGCAAAGTCCCTGAC GGGCACGGCCTC-3'	+893
	FUT2 P16 (R)	5'-CTCCCTGTGCCTTGGAAAGTAT-3'	+1094
	FUT2 P17 (F)	5'-AACTGCACTGCCAGCTTCATGC-3'	-83

15 **Table 2:** Overall Recombination Fractions ( $\theta$ ), Lodscores (Z) And Number Of Informative Animals (N) For *M307* And Loci Of The *HAL* Linkage Group In The Landrace Experimental Population

	Locus pair	N	$\theta$	Z
20	<i>S-ECF18R</i>	183	0.01	50.6
	<i>M307-S</i>	183	0.01	50.6
	<i>M307-ECF18R</i>	216	0.01	57.1
	<i>M307-RYR1</i>	198	0.02	47.2
	<i>M307-GP1</i>	147	0.03	34.2
25	<i>M307-PGD</i>	147	0.04	24.5

<sup>2</sup> Primers *FUT1* P10 and *FUT1* P11 are derived from the human *FUT1* gene.

**Table 3:** Haplotype Frequencies At the Four Loci (*S-FUT-1* (*M307*, *M587*)-*ECF18R-RYR1*) In the Landrace (SL) Experimental Population And Randomly Selected Large White (LW) Pigs.

5	Breed	Haplotype <sup>3</sup> at <i>S</i> , <i>FUT1</i> ( <i>M307</i> , <i>M587</i> ), <i>ECF18R</i> , <i>RYR1</i>	Frequency <sup>4</sup> (number)
	SL	<i>sAGbT</i>	70 (28)
		<i>sAGbC</i>	5 (2)
		<i>sGGBC</i>	15 (6)
		<i>sGABC</i>	10 (4)
	LW	<i>sAGbC</i>	56 (9)
		<i>sGGBC</i>	31 (5)
		<i>sGGBC</i>	13 (2)

3

*S*: Suppressor locus for A and O blood types (*S* and *s*).  
*FUT1* (*M307*): alteration of adenine (A) to guanine (G) at nucleotide 307 of the alpha (1,2) fucosyltransferase (*FUT1*) gene. *FUT1* (*M587*): alteration of adenine (A) to guanine (G) at nucleotide 857 of the *FUT1* gene. *ECF18R*: *E. coli* F18 receptor. The dominant susceptible allele is indicated by *B* and the resistant allele by *b*. *RYR1*: skeletal muscle ryanodine receptor. C (cytosine) is the dominant resistant and T (thymine) the susceptible allele for malignant hyperthermia.

<sup>4</sup>Haplotype frequencies in % and absolute number of haplotypes between brackets.



**Table 4:** Distribution Of The Genotypes, Tetrachoric Correlation (R) And Significance Of The Association ( $X^2$  And  $w \times X^2$ ) Of The Associated Polymorphic *FUT1* (*M307*) And *ECF18R* Loci In Landrace (SL) Experimental Population And Randomly Selected Large White (LW) Swine.

	Breed	Locus	<i>FUT1/M307</i>			r	X <sup>2</sup>	X <sup>2</sup> x w <sup>5</sup>
			Genotype	A/G	A/A			
10	SL	<i>ECF18R<sup>c</sup></i>	<i>b/b</i>	1	113	0.98	213.1	42.6***
			<i>B/B</i>	106	1			
			<i>A/G</i>					
			Genotype ( <i>A/G</i> )	<i>G/G</i>	<i>A/A</i>			
	LW	<i>ECF19R<sup>d</sup></i>	<i>b/b</i>	0	5	1.00	29.0	11.6***
15			<i>B/b, B/B</i>	24	0			

## METHODS

### 1. Primers

Primers derived from the human *FUT1* gene were used for the amplification of its porcine counterpart from genomic DNA. From the resulting porcine sequences specific primers were designed which were used in further amplification and sequencing reactions (Table 1).

### 2. Screening of a Porcine Genomic Library

Porcine genomic libraries were screened with either a porcine *FUT1* probe obtained with primers P7 and p10 or a porcine *FUT1* cDNA. A porcine genomic library, constructed in SuperCos 1 (Stratagene, LaJolla, Ca, USA), was screened with an  $\alpha^{32}P$  dATP labeled (Prime It II, Stratagene) *FUT1* probe obtained from porcine genomic DNA with primers P7 and P10. After hybridization of replica filters at 42°C for 15h (50% formamide, 6 x SSC, 5 x Denhardt's, 0.5% SDS, 0.1 mg/ml Salmon Sperm) and washing

5

A weight factor of  $w=0.2$  (SL) and 0.4 (LW) was applied to correct for the lack of precision resulting from inclusion of related animals in the data, according to Cotterman (1947). \*\*\* $p > 0.001$ .

6

Animals of genotype *b/b* at the *ECF18R* locus are resistant and those of genotype *B/b* and *B/B* are susceptible to adhesion of F18ab *E. coli* bacteria.

twice at 65°C for 30 min. (1x SSC, 0.1%SDS), positive colonies were identified after exposure (15h, -80°C) to X-ray film.

### 3. In Situ Hybridization of Porcine Metaphase Chromosomes

Cosmid clones ETHs1, ETHs2, ETHs3, ETHs4 and ETHs6 were subjected to fluorescence *in situ* hybridization (FISH) (Solinas Toldo, *et al.*, 1993) or direct *in situ* chromosomal PCT (DISC PCR) on porcine metaphases. Metaphase chromosomes were Q-banded and photographed before hybridization. The probes were labeled by random priming using biotin-16-dUTP. Signal detection and amplification was performed using avidin-FITC and biotinylated anti-avidin. The chromosomes were counterstained with 4,6-diamidino-2-phenylindole, and the relative positions of the cosmids were determined as described by Solinas Toldo, 1993.

### 4. Subcloning

Enzymatic digests of probe positive genomic colonies were separated on agarose gel, transferred to a nylon membrane, and probe positive bands were subcloned into plasmids for *FUT1* sequencing. The sequence of *FUT1* derived from this method is shown in FIG. 1.

KspI-, EcoRI- and KspI/EcoRI digests of all cosmids were separated on a 0.8% agarose gel and transferred to a Hybond N nylon membrane. (Meijerink *et al.*, 1997). This blot was hybridized with  $\alpha^{32}\text{P}$  dATP labeled porcine *FUT1* PCR products (primers P6-P11 and P7-P10). Based on the autoradiographic signals, ETHs1, -s2 and -s3 were subjected to further subcloning into pBluescript SK - (Stratagene), and *FUT* sequences were determined from subclones. The sequences of two *FUT*-like open reading frames (ORFs) (*FUT1* and *FUT2*) obtained from cosmids ETHs2 and -s3 were compared in *ECF18R* positive (*BB/Bb*) and negative (*bb*) animals by direct sequencing of PCR products.

### 5. Polymerase Chain Reaction and Direct Sequencing

Using the Perkin Elmer Ready Reaction Dye Terminator kit (Perkin Elmer Cetus, Norwalk, CT, USA) and 10 pmol of primer, cycle sequencing was performed with a thermal program consisting of an initial denaturation of 5 min at 95°C, followed by 25 cycles of 30 sec 95°C, 15 sec 50°C and 4 min 60°C. Primers used for amplification and sequencing of the porcine alpha (1,2) fucosyltransferase genes are listed in Table 1. Additional primers were designed taking the possibility of cross-annealing of primers due to the high similarity of *FUT1*, *FUT2* and the *FUT2* pseudogene into account. Samples

were analyzed on a 373A ABI sequencer (Applied Biosystems, Inc.) and sequence analysis was performed with the GCG package (Devereux, 1984).

6. Production of Informative Offspring

Single nucleotide polymorphisms were analyzed in 221 Landrace swine produced from 4 boars and 16 sows, and in 29 Large White swine produced from 9 matings between unrelated swine. In order to produce a large number of informative offspring for the examination of linkage between porcine genes encoding ECF18 receptors and selected polymorphic loci, only informative Landrace matings of the type *B/b* x *b/b* were produced.

7. Colonization Test

In a study of Bertschinger *et al.*, 1993, the above mentioned Landrace swine were also tested for ECF18 susceptibility in a colonization test. For this, swine were inoculated shortly after weaning with bacteria of *E. coli* strain 124/76 of serotype 0139:K12(B):H1:F18. (Rippinger, *et al.*, 1995). Faecal shedding of the bacteria was monitored daily. The extent of colonization was calculated as the mean of the two highest faecal scores. Swine with a mean faecal score of 3.5, corresponding to 6.7 log colony forming units (CFU)/g or more, were considered susceptible to colonization. This limit was based on a lack of mortality below this value, and on scores obtained from completely resistant litters.

8. Linkage Analysis of Nucleotide Polymorphisms

The results of the single nucleotide polymorphisms were compared with typing data for ECF18R, which were identified in an *in vitro* adhesion assay described by Vögeli *et al.*, (1996), and with typing data for the *GPI*-, *PGD*-,  $\alpha$ -1-B-glycoprotein-(AIBG), ryanodine receptor (*RYR*), *EAH*- and *S*-loci as published by Vögeli *et al.*, (1996). Pairwise linkage analysis and calculation of recombination fractions was performed using the CRI-MAP version 2.4 programme (Green *et al.*, 1990). Multipoint linkage analysis was performed by sequential insertion of the above loci into the map. Haplotype frequencies were calculated from the parental animals, in the Landrace families and from the 8 parental Large White animals which were haplotyped for from progeny information. Tetrachoric correlations of *ECF18R* and mutations in *FUT1* (*FUT1/M307*) (polymorphisms) were calculated on all Landrace and Large White progeny.

9. Southern Blot Analysis

Southern blot analysis was performed on cosmids ETHs (1-3), ETHs2 (4-6) and ETHs3 (7-9) after digestion with enzymes KspI (1,4,7), EcoRI (2, 5, 8) and KspI/EcoRI (3, 6, 9) and separation on 0.8% agarose. Hybridization with an  $\alpha^{32}$  PdATP labeled 5'

- 5 *FUT1* fragment (primers P6-P11) results in the same hybridizing 940 bp band in both the KspI digest (lane 4) and the KspI/EcoRI digest (lane 6). However, hybridization with a 3' *FUT1* fragment (primers P7-P10) (Table 1) shows a 6.2 kb KspI band in lane 4 and a 1.1 kb KspI/EcoRI band in lane 6. Both the 5' and 3' *FUT1* fragments hybridize to the same 4.6 kb EcoRI fragment in lane 5. This indicates the presence of a KspI site in the
- 10 *FUT1* gene contained in cosmid ETHs2. Cross hybridization of the 3' *FUT1* fragment detects 2.7 kb (lanes 2, 3, 8 and 9) and 8.2 kb (lanes 8 and 9) bands, resulting in the identification of the *FUT2* pseudogene (incomplete ORF) and the *FUT2* gene sequences, respectively.

10. Restriction Fragment Polymorphisms

- 15 Detection of (A) the M307 G to A and (B) the M857 G to A mutation in the porcine *FUT1* gene was achieved by restriction length polymorphism analysis, using various restriction enzymes. Digestion of amplified *FUT1* fragments with CfoI (A) and AclI (B) results in a restriction fragment polymorphism. In the first lane is a 100 bp marker. Fragment lengths are indicated in base pairs. (A) The *M307<sup>ΔA</sup>* genotype (lane
- 20 2) generates 328 and 93 bp restriction fragments while the *M307<sup>G/G</sup>* genotype (lane 4) generates 93, 241 and 87 bp fragments and heterozygous *M307<sup>ΔA/G</sup>* genotypes (lane 3) shows all four fragments.

(B) Digestion of the *M857<sup>ΔA</sup>* genotype (lane 2) generates 174 bp fragments, while it generates 136 and 38 bp fragments in the *M857<sup>G/G</sup>* genotypes (lane 4), and in *M857<sup>ΔG</sup>* genotypes (lane 3) all three fragments are generated.

11. Source of Swine

Data of the Swiss Landrace experimental population came from two pedigrees, which were built up at the Institute of Veterinary Bacteriology, University of Zurich. All other pigs of the Large White, Swiss Landrace, Duroc, Hampshire and Pietrain breeds

30 came from different breeding herds of Switzerland. Other swine were randomly obtained from farms in the U.S. Midwest.

**DOCUMENTS CITED**

- Bertschinger *et al.* (1993) *Veterinary Microbiology* 35:79-89.
- Cohney *et al.* (1996) *Immunogenetics* 44:76-79.
- Devereux *et al.* (1984) *Nucleic Acids Res.* 1:387-395.
- 5 Fujii *et al.* (1991) *Science* 253:448-451.
- Green *et al.* (1990) Documentation for CRI-MAP, Version 2.4, St. Louis:  
Washington University School of Medicine.
- Kelly *et al.* (1994), *Proc. Natl. Acad. Sci., U.S.A.* 91:5843-5847.
- Meijerink, E. *et al.* (1997), 25<sup>th</sup> Int. Conf. on Animal Genetics, p. 44.
- 10 Rippinger, *et al.* (1995) *Vet. Microbial.* 45:281-295.
- Solinas Toldo *et al.* (1993) *Mamm. Genome* 4:720-727.
- Thurin and Blaszczyk-Thurin, (1995) *J. Biol. Chem.* 270(44):26577-26580.
- Vögeli *et al.* (1996) *Animal Genetics* 27:321-328.
- Vögeli *et al.* (1997) *Schweiz. Arch. Tierheilk.* 139:479-484.
- 15 U.S. Pat. No. 5,358,649, MacIennnon *et al.*
- U.S. Pat. No. 5,552,144, Valery *et al.*
- WO 8604604 987P, Inventor: Peterson.
- WO 9628967, Inventor: Koike, C.
- WO 9413811, Inventors: Imberechts and Lintermans.
- 20 TW 266264, Inventors: Jeng and Liou.